Claims

We claim:

1. A computer-implemented method for mapping a user data schema to a mining model schema, comprising:

matching columns of the user data schema to corresponding columns of the mining model schema to provide a mapping;

determining whether data within matching columns of the user data schema has a data type different than data within the corresponding columns of the mining model schema;

transforming the data within the matching columns of the user data schema if the data type is determined to be different; and

updating a matching resource based on the mapping.

2. The method of claim 1, further comprising:

providing an opportunity to manually alter the mapping after transforming the data; and

presenting a final view of the mapping after providing the opportunity, wherein the updating step is performed after the final view is presented.

3. The method of claim 1, wherein the matching step comprises determining whether names of the columns of the user data schema exactly match names of the columns of the mining model data schema.

4. The method of claim 3, wherein the matching step further comprises determining whether the names of the columns of the user data schema are similar to the names of the columns of the mining model data schema based on the matching resource.

5. The method of claim 4, wherein the matching step comprises determining whether the names of the columns of the user data schema match the names of the columns of the mining model schema based on one or more formulae.

6. The method of claim 5, wherein the matching step further comprises determining whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema.

7. The method of claim 1, wherein the matching resource is selected from the group consisting of a thesaurus, a dictionary and a similarity threshold.

8. The method of claim 1, further comprising:

populating a schema consolidation table with names of the columns of the mining model schema, prior to the matching step; and

updating the schema consolidation table with names of the matching columns of the user data schema, during the updating step.

9. A computer-implemented method for mapping a user data schema to a mining model schema, comprising:

populating a schema consolidation table with names of columns of the mining model schema;

mapping the user data schema to the mining model schema by matching columns of the user data schema to corresponding columns of the mining model schema;

determining whether data within matching columns of the user data schema has a data type different than data within the corresponding columns of the mining model schema;

transforming the data within the matching columns of the user data schema if the data type is determined to be different;

providing an opportunity to manually alter the mapping after transforming the data;

presenting a final view of the mapping after providing the opportunity to manually alter the mapping; and

updating a matching resource and the schema consolidation table based on the mapping.

10. The method of claim 9, wherein the matching step comprises determining whether names of the columns of the user data schema exactly match names of the columns of the mining model schema.

11. The method of claim 10, wherein the matching step further comprises determining whether the names of the columns of the user data schema are similar to the names of the columns of the mining model data schema based on the matching resource.

12. The method of claim 11, wherein the matching step comprises determining whether the names of the columns of the user data schema match the names of the columns of the mining model schema based on one or more formulae.

13. The method of claim 12, wherein the matching step further comprises determining whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema.

14. The method of claim 9, wherein the matching resource is selected from the group consisting of a thesaurus, a dictionary and a similarity threshold.

15. The method of claim 9, wherein the step of updating the schema consolidation table comprises updating the schema consolidation table with names of the matching columns of the user data schema.

16. A computerized system for mapping a user data schema to a mining model schema, comprising:

a column matching system for matching columns of the user data schema to corresponding columns of the mining model schema to provide a mapping;

a model differentiation system for determining whether data within matching columns of the user data schema has a data type different than data within the corresponding columns of the mining model schema;

a data transformation system for transforming the data within the matching columns of the user data schema if the data type is determined to be different; and

an update system for updating a matching resource based on the mapping.

17. The system of claim 16, further comprising:

a manual matching system for providing an opportunity to manually alter the mapping after transforming the data; and

a view system for presenting a final view of the mapping after providing the opportunity.

18. The system of claim 16, wherein the column matching system determines whether names of the columns of the user data schema exactly match names of the columns of the mining model data schema.

19. The system of claim 18, wherein the column matching system further determines whether the names of the columns of the user data schema are similar to the names of the columns of the mining model data schema based on the matching resource.

20. The system of claim 19, wherein the column matching system further determines whether the names of the columns of the user data schema match the names of the columns of the mining model schema based on one or more formulae.

21. The system of claim 20, wherein the column matching system further determines whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema.

22. The system of claim 16, wherein the matching resource is selected from the group consisting of a thesaurus, a dictionary and a similarity threshold.

23. The system of claim 16, further comprising a table population system for populating a schema consolidation table with names of the columns of the mining model schema, wherein the update system further updates the schema consolidation table with names of the matching columns of the user data schema.

24. A program product stored on a recordable medium for mapping a user data schema to a mining model schema, which when executed, comprises:

program code for matching columns of the user data schema to corresponding columns of the mining model schema to provide a mapping;

program code for determining whether data within matching columns of the user data schema has a data type different than data within the corresponding columns of the mining model schema;

program code for transforming the data within the matching columns of the user data schema if the data type is determined to be different; and

program code for updating a matching resource based on the mapping.

25. The program product of claim 24, further comprising:

program code for providing an opportunity to manually alter the mapping after transforming the data; and

program code for presenting a final view of the mapping after providing the opportunity.

26. The program product of claim 24, wherein the program code for matching determines whether names of the columns of the user data schema exactly match names of the columns of the mining model data schema.

27. The program product of claim 26, wherein the program code for matching further determines whether the names of the columns of the user data schema are similar to the names of the columns of the mining model data schema based on the matching resource.

28. The program product of claim 27, wherein the column matching system further determines whether the names of the columns of the user data schema are similar to the names of the columns of the mining model schema based on one or more formulae.

29. The program product of claim 28, wherein the program code for matching further determines whether the data within the columns of the user data schema corresponds to the data within the columns of the mining model data schema.

30. The program product of claim 24, wherein the matching resource is selected from the group consisting of a thesaurus, a dictionary and a similarity threshold.

31. The program product of claim 24, further comprising a program code for populating a schema consolidation table with names of the columns of the mining model schema, wherein the program code for updating further updates the schema consolidation table with names of the matching columns of the user data schema.